Critical Challenges to Gamifying Education
A Review of Central Concepts

J. Tuomas Harviainen, Ph.D.
School of Information Sciences
University of Tampere, Finland
jiituomas-at-gmail-com

This paper discussed key problems in the deployment and assessment of game-based learning and gamified educational systems. Utilizing a metatheoretical survey of prominent references, as well as analytic summaries of empirical research, it shows that several challenges to learning through game-based systems still exist, and that these problems have to a large extent been identified decades ago. These include excessive competition, cheating, harmful stealth learning, a false sense of safety, lack of accurate evaluation, the ways in which game-based learning and gamification are reliant on personal and contextual factors, and the treatment of theoretical works as implicitly reliable by many researchers. As a result, the paper suggests a move towards emphasizing actual research and empirics, and the balanced optimization of the levels of fun and challenge during play.

Introduction

In this paper, I present points of critique that are necessary to take into account when applying techniques of digital game-based learning (GBL) and the gamification (turning activities into gameful experiences by adding motivational affordances; Huotari & Hamari, 2012) of educational tasks. Systematic game-based learning has been around for over two centuries, even before the advent of von Reisswitz' *Kriegspiel* in 1812 (Meier, Newell & Pazer, 1969), and the history of training games and simulations is at least 3,000 years old (Keys & Wolfe, 1990). However, we still know relatively little of how and why games function as educational tools, even with the existence of over half a decade of research literature (ranging from e.g., Kibbee, Craft & Nanus, 1961, to Whitton, 2014), several specific journals such as *Simulation & Gaming* (est. 1970) and *International Journal of Game-Based Learning* (est. 2011), and numerous successful deployments of games in various learning contexts. We know even less about which parts of gamification actually work. This problem has nevertheless not, in the least, prevented the use of such techniques, nor the sales of them by marketeers as something that they present as automatically superior to traditional methods of both teaching and organizational development.

Based on a metatheoretical survey of existing critical and empirical research, I discuss the ways in which the goals of using game techniques for allotelic (external to the activity itself) purposes may get sidetracked, as well as why. This is constructed as a meta-analysis of earlier works, particularly those that contain empirical evidence from more than just a singular case. These examples are presented in reflection with data gathered on game-based learning over the last six decades, to point out that this critique is not new. On the contrary, it has been discussed alongside game-based teaching techniques from early on, and only recently forgotten and then re-discovered. I believe that the reason for this is that the new wave of GBL was not inspired by existing traditions of using games and simulations for educational purposes, but rather the increasing popularity of commercial digital games. This is exemplified by books such as *Serious Games: Mechanisms and Effects* (Ritterfeld, Cody & Vorderer, eds., 2009), in which only two articles out of 28 take into account the earlier tradition of serious games research in any way.

The following chapters assess problems with citation chains, player relations to games and simulations, assessment, safety to fail and, most significantly of all, what the players learn. Alongside these questions is the issue of whether gamification, which seems to rely on context and individual variables in addition to its mechanisms for positive effects (Hamari, Koivisto & Sarsa, 2014), actually fits educational purposes. The

research question of this paper is whether we currently know enough about gamifying education to safely and beneficially utilize that gamification.

An Untied Safety Rope

Game-based learning depends on the use of a game to support learning goals in a manner where it is a tool and not a distraction (Whitton, 2009). Gamification seeks to invoke the same kind of psychological experiences as do games (Huotari & Hamari, 2012), and its affordances need to be the same as in games (Deterding et al., 2011). Therefore, successful deployment of such techniques and technologies is an informed task that requires expertise. Without sufficiently accessing earlier research in a broad spectrum, any gamification of learning faces the problem of a chain of references at the end of which no empirics can be found. At their core, many works on gamification and/or GBL rely on texts that were simultaneously non-empirical and highly ideological. A handful, in particular, stand out James Paul Gee, in his seminal *What Video Games have to Teach us About Learning and Literacy* (2007, originally published in 2003), presents a set of conjecture on how games *might* improve or facilitate learning, based on an analysis of 36 video games. No empiric evidence is included, and Gee himself effectively acknowledges this (ibid.). Those who cite him, however, often do not. Instead, his statements are presented as facts and not theory.

In a similar fashion, Jane McGonigal' *Reality is Broken* (2011), a central (but far from the first) book on gamification, contains the author's opinions on how adding game elements to other activities may function, yet ends up cited as factual. Furthermore, certain popular works on gamification techniques, produced by people whose training and experience is in marketing and not education, may actually contradict existing research and accumulated data   (see Zichermann & Cunningham, 2011, for a key example, and Deterding, 2011, for critique relating to it). For example, it appears that a large segment of players do not play for the sake of achievements (Warmelink, 2014), yet gamification marketeers claim the opposite (see e.g., Zichermann & Linder, 2010). Using such unconfirmed sources is therefore very risky, as applying such material to practice without sufficient criticality and pre-knowledge may turn out harmful to the intended learning experiences. One tool may not fit another context, and it appears that successful gamification is highly context-dependant (Hamari, Koivisto & Sarsa, 2014). Therefore, uncritical transportation of such concepts is problematic. This has also been confirmed from the side of organizational learning: people who prefer their online organizations flexible and playful may not want the same from their workplaces (Warmelink, 2014). The foundational practices differ between gaming groups, as they do between other organizations (see Vesa, 2013), but it is probable that gamelike practices will be increasingly seeping into everyday life. This will render some gamifications mundane (Warmelink, 2014), and therefore likely take away their motivation-increasing novelty as well.

The result of these problems is that a significant error margin exists, unless one is able to tailor the game to the subject matter being taught, and make sure it stays on course. If the tailoring, in turn, is based on non-empirical, ideologically created principles, the risk reappears. It nevertheless bears noting that some of the early conjecture has in fact been proven right by later researchers, with actual empirics. Therefore, it is recommendable that before deploying tools based on shakily grounded ideas - or dismissing the works critiqued above - one at least look also at the works that cite those within a similar area of interest as the intended game or gamification deployment.To accomplish that, we now turn to look at the challenges and limitations already identified by earlier researchers.

Framing-related Challenges

As noted by e.g., Whitton (2009) and Warmelink (2014), not every person attending a classroom is a digital native, no matter what Prensky (2001) may claim. Individual variance on attitudes, aptitudes and access to technology still exists. As games tend to rely on existing conventions to support their play (e.g., the interface of a new real-time strategy game tends to resemble preceding ones), for easier access (Jørgensen, 2013), the importance of earlier knowledge comes into play. Games do not function at all without players' brought-with pre-knowledge (Crookall, Oxford & Saunders, 1987). Furthermore, not everyone likes games, or wants to play (Whitton, 2009; Harviainen & Savonsaari, 2013), and an unwilling player may disrupt the game for all involved.

Of the ones who do like games, a certain subset is also a problem group: those who are highly competitive. Many players of educational games see all games as competitions, but this particular group is so driven that they may miss the learning goals almost entirely (Harviainen, Lainema & Saarinen, 2014). Ian Bogost (2007), following Gee (2007), states that in order to learn from a game, the player needs to apply his or her procedural literacy - the ability to read and analyze the game as a system and to compare it to other systems, including the physical world. Highly competitive players, however, utilize their procedural literacies to outperform others, yet do so using the logic of that system and not by applying the learning task at hand. This can lead them to learn completely wrong things (Harviainen, Lainema & Saarinen, 2014). While trying to beat the game, they do not observe the variables of that game (Carlson & Misshauk, 1972). Highly competitive framing is furthermore a problem not just those particular players: by emphasizing the competitive aspect too much, they create a sense of loss in the other participants, too. That sense may overtake the learning goals and make those who do not perform well enough to feel like failures (Graham & Gray, 1969; Zichermann & Linder, 2010). A similar effect has been observed of gamified education: it too seems to increase competition in a sometimes problematic manner (Hakulinen, Auvinen & Korhonen, 2013).

Competitive framing also encourages cheating, which arises partially from the supposedly safe environment of play where consequences do not carry weight (see e.g., Tsuchiya & Tsuchiya, 1999), and the desire to win. This is not only an assessment problem, but again also that of whether the correct things are learned while playing. Cheating is a distraction to learning (Henriksen, 2008). It can be argued that a player able to cheat is exhibiting particularly good procedural literacy relating to that particular game - and is thus showing more learning - but it may just as well be the use of an accidentally discovered bug in the program, meaning nothing is learned (Harviainen, Lainema & Saarinen, 2014). Again, this impacts other players as well: if knowledge of someone cheating at an educational game spreads, the anchoring of the learning goals will be made more difficult for all (or nearly all) participants, as the cheating reduces the value of achievement in that game for everyone.

This ties into the issue of the game not "being real". While players expect to be able to ignore the consequences of their play to some extent, not all consequences are neutralized. They can, for example, in certain games at the last minute sell all their production machinery, or all workers into slavery (Harviainen, Lainema & Saarinen, 2014). These are not real consequences, as the company and the workers are not real. However, they may nevertheless face repercussions, in the form of performance assessment, but also through social feedback. Graham and Gray (1969) use the example of how it may be risky to outperform one's boss in a simulation, but a similar case can be made of e.g., triumphing over the class bully, or of simply being too good a player in a social group where outshining is disliked. Because of this, participants may sometimes play far below their actual capability, and as a result, learn much less than they would in the context of more traditional methods.

Learning Goals and Challenge

Just as not everyone likes games, neither do they fit everyone's learning style (Whitton, 2014). While the added, playful motivation may assist weaker learners to do better, and make the average student at least enjoy the learning more (and often also perform better), less is known about particularly talented students. For some, traditional cramming will be the optimal learning style, and any gamification of the task simply a distraction. Others may simply prefer a learning method other than games (see e.g., Prensky, 2001).

The harsh fact is that we still know very little about how and what exactly people learn from playing games, even specifically designed educational games (Klabbers, 2003). According to Prensky (2001), players supposedly learn simply by playing the game, without themselves often even realizing it. Whitton (2009), as well as Harviainen, Lainema and Saarinen (2014), however disagree: without sufficient reflection during and after play, such "stealth learning" produces faulty, unreliable results. Furthermore, without a debriefing to anchor the learning, even if the intended things were learned, they are likely to fade away (Kim, 1993).

To promote its intended learning goal, a game has to be geared for that task, either through design, pre-play

briefing, or both. The content has to be presented in a logical learning sequence (Graham & Gray, 1969). Optimally, it should also be iterative, so that the learning builds upon earlier experiences in play and adds more learning content each time, while working also as a memorization system for what has already been learned while playing (Henriksen, 2008). The level of challenge has to match player skill sufficiently, so that it is neither too easy nor too hard (Carlson & Misshauk,1972; Henriksen, 2008). Deviating from that zone risks both the educational goal and the player enjoyment. Likewise, too much fun can also be problematic: if the player is too engaged in the play, procedural literacy is applied without sufficient reflection, and the learning goals are lost (Henriksen, 2008; see also Bogost, 2007). Without sufficient fun, however, the activity becomes tedious and the iterative content simply boring repetition (Graham & Gray, 1969). This also happens when the novelty of gamification wears off (Hamari, 2013).

As a result, successful deployment of games and game elements as learning tools is first and foremost an issue of checks and balances, between educational content and playability, between too little and too much challenge, between insufficient and excessive enjoyment - and between a very wide range of personal and interpersonal variables affecting those. Yet even if one balances all those correctly, the goal may not be fully obtained, because well-learned does not necessarily mean easily assessed, and the educational establishments require proper assessment.

Assessment

It is tempting to use performance at play as the key method of grading the use of the game. However, the player who gets the highest score may not be the one who learns the most (Harviainen, Lainema & Saarinen, 2014). As described above, high-score players may likewise have engaged in practices directly harmful to their own learning, as well as to that of others. "Performance does not equate learning" is a principle that has to be applied to the assessment process (Knotts & Keys, 1997). It can certainly be a part of the grading process nevertheless,just not all of it.

The deployment of a game in an educational context reflects more than just the needs of the curriculum. While gameplay does allow players some leeway to act freely within the constraints of the game, the subject matter is not treated in a truly neutral fashion. It always contains values embedded in the game by its designer, as well as the value sets of the teacher who has chosen to use that particular game for that particular topic (Graham & Gray, 1969). This is no different from a course book and the teacher's choice to use it. Because of the educational framework, and the player desire to do well (often: win), players have limited chances and/or motivation to explore alternative options (Carlson & Misshauk, 1972). Thus, they have less room for reflection. Likewise, they have few chances to engage in what Myers (2010) calls "functional bad play" - the exploration of the system by playing intentionally against designer expectations. Such explorations of variables, however, are necessary for more holistic learning in a simulation or game setting, and actually one of those settings' key advantages (Crookall, Oxford & Saunders, 1987). This includes the ability to try out strategies that fail, intentionally or not, something that is incompatible with excessive competitiveness and cheating, and also with a very performance-based evaluation. Those players who fail the most may actually be the ones learning the most (Lainema, 2003).

Due to these factors, alternative means of learning assessment are required. Harviainen, Lainema and Saarinen (2014) recommend the use of reflective essays that are tied into the content of the course, as a part of which the game is being deployed. Whitton (2009) suggests the use of creative tasks, such as portfolios, reports, discussion posts or reflective accounts, as the basis of the grading. She also sees in this an opportunity for the learners to be creative and to adapt their game-learned skills into the real world. The assessment process itself becomes a part of the reflection and anchoring.

Players can during play also perform some activities that demonstrate learning, yet do not fit the pre-determined systems for assessment. For example, Palmunen et al. (2013) report how players of a business simulation purchased all materials of a certain type, so as to deprive other players of the chance to produce their products. This is something, they note (ibid.), that falls outside traditional assessment tools such as Bloom's (1956) taxonomy, yet clearly demonstrates the kind of procedural literacy described by Bogost

(2007). The teacher therefore has to examine also the play itself, as it may reveal learning that the basic assessment system does not.

Simultaneously to assessing student learning, the deployed game or gamification technique itself also requires evaluation. As pointed out above, games are not suitable teaching tools for all occasions, students, or subjects. Their challenge level may not be optimal for the task, nor the engaging fun. No game, even the most simulation-oriented one, is perfect. This is both an issue of convenience and a requirement: the fact that games are functional but limited representations is what makes them playable and interesting (Klabbers, 2009). A side effect of this is that the more complex and perfect the system - i.e., the more elements there are to potentially learn and assess - the less we can trust players to master during the course of a single deployment. As observed by Palmunen et al. (2013), novices rarely have the the skills needed to compare the system's full extent with the real world, and thus they cannot really apply their procedural literacies the way a teacher might expect.

Therefore, each deployment of a learning game is an evaluation of both the students and the game itself (and, if the teachers are honest to themselves, of the deploying teachers as well). Connolly et al. (2008) suggest seven criteria for this: learner performance, student motivation, perceptions, attitudes towards the subject and towards learning that subject, preferences in learning style, factors associated with the game itself as a learning environment, and collaboration. Connolly et al. (ibid.) see the last of these as optional and dependent on the game's design, whereas Whitton (2009) points out collaboration's importance to learning design, particularly for adult learners.

Finally, assessment also needs to take into account the students' experience of the game: how they themselves perceived the challenges, engagement level, interaction opportunities and so forth. For this task, too, the reflective approach - whether in the form of essays or other creative forms - is particularly suitable (Harviainen, Lainema & Saarinen, 2014; Whitton, 2009).

Discussion

As was already observed in the 1960's, rewards and satisfaction from play are highly individual things, varying from one person to the next. Likewise, motivation for learning is highly individualistic (Graham & Gray, 1969). The game-frame alone is insufficient to function as something that would be the right kind of motivation for all players (Prensky, 2001; Whitton, 2009). Gamified systems furthermore depend on context, the personas of the participants, and environmental factors such as time available (without which increasing e.g., time-on-task is impossible). As shown by Hamari, Koivisto and Sarsa (2014), participant reactions to gamification seem to be generally more positive than negative, but the very same elements that the majority appears to find positive are the ones a smaller segment of participants dislikes.

One often overlooked aspect of gamifying education is that it may make identifying and evaluating the learning task difficult (Dominguez et al., 2013). Therefore, just like any game deployed for learning purposes, it requires a pre-play briefing and a thorough debriefing. Without these, participant motivation will be lower, and both the intended, overt learning and (especially) any stealth learning will be at an increased risk of ending up soon forgotten, or teaching the wrong thing. Gamification deployment in education has to learn from the deployment of games for learning and the decades of research gathered for that purpose. The research should furthermore be of an empirical kind, so as to avoid confusing marketing hype or educational speculation with actual research. At the same time, however, the non-grounded work too has its place, as inspiration for further research and testing.

The gamification of educational tasks - whether done by adding game elements or by deploying games - is first and foremost a question of seeking informed balances. Done correctly, existing research suggests, the results are positive and increase participant motivation, thus making learning more probable - at least for many participants. Done wrong, such play is at best wasted time, time that could have been better utilized for other types of learning techniques.

One final thing that needs to be mentioned is that when using game-based or gamified methods, educators should take time during briefing and debriefing to discuss with the students the ways in which the game is different from the real world. This is partially for the purpose of improving the students' information literacies, but it has a deeper purpose: the fields in which educational games have been used the longest - military and business sciences - are also those that are especially associated with treating real-world phenomena (human lives and entire national economies included) as if they were games, as well as getting away with doing so. This is an extremely risky, unwelcome type of gamification, the support of which through an educational system is not recommended.

Conclusions

Current research appears to support the hypothesis that in general, games and gamification do function as useful, positive tools for learning. At the same time, however, the very elements that motivate some learners (e.g., competition) are the ones disliked by others, and also those that when taken to excess, cause problems. We still do not really know how people learn with and through games and gamified systems. More empirics are needed, as is the further study of the contextual dependence that seems to affect what is learned through gameplay. All the evidence points to the direction that games and gamification do offer new potential for improved learning, at the very least for some people. Yet, to produce positive results, the design and deployment of such tools requires informed skill, not just hopeful thought and black-box thinking. These are learning practices we must apply as the careful scientists that we are, and not as magicians who trust that if we just follow the formula, fixed results will always appear.

References

Bloom, B. S. (1956). *Taxonomy of educational objectives, handbook 1: The cognitive domain*. New York: David McKay Co. Inc.

Bogost, I. (2007) *Persuasive games: The expressive power of video games*. Cambridge, MA: The MIT Press.

Carlson, J. G. H & Misshauk, M. J. (1972). *Introduction to gaming: Management decision simulations*. New York: Wiley.

Crookall. D., Oxford, R. & Saunders, D. (1987). Towards a reconceptualization of simulation: From representation to reality. *Simulation/Games for Learning,* 17(4), 147-171.

Deterding, S. (2011). A quick buck by copy and paste. A review of "Gamification by design". Blog post with commentary. Retrieved May 28, 2014, from http://gamification-research.org/2011/09/a-quick-buck-by-copy-and-paste/

Deterding, S., Dixon, D., Khaled, R. & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, Tampere, Finland, September 28-30, 2011, ACM, pp. 9-15.

Dominguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernández-Sanz, L., Pagés, C. & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63, 380-392.

Gee, J. P. (2007). *What video games have to teach us about learning and literacy*, revised and updated edition.New York: Palgrave Macmillan.

Graham, R. G. & Gray, C. F. (1969). *Business games handbook.* United States of America: American Management Association.

Hakulinen, L., Auvinen, T. & Korhonen, A. (2013). Empirical study on the effect of achievement badges in

TRAKLA2 online learning environment. In *Proceedings of Learning and Teaching in Computing and Engineering (LaTiCE) Conference*, March 21-24, 2013, Macau, pp.47-54.

Hamari, J. (2013). Transforming Homo Economicus into Homo Ludens: A field experiment on gamification in a utilitarian peer-to-peer trading service. *Electronic Commerce Research and Applications*, 12(4), 236-245.

Hamari, J., Koivisto, J. & Sarsa, H. (2014). Does gamification work? - A literature review of empirical studies on gamification.   In *Proceedings of the 47th Hawaii International Conference on System Sciences*, Hawaii, USA, January 6-9, 2014.

Harviainen, J. T., Lainema, T. & Saarinen, E. (2014) Player-reported impediments to game-based learning. *Transactions of the Digital Games Research Association*, 1(2), 55-83.

Harviainen, J. T. & Savonsaari, R. (2013). Larps in high schools. In Moseley, A. & Whitton, N. (Eds.) *New traditional games for learning*, pp 134-145. London: Routledge.

Henriksen, T. D. (2008). Extending experiences of learning games - or why learning games should be neither fun, educational or realistic. In Leino, O., Wirman, H. & Fernandez, A. (Eds.) *Extending experiences: Structure, analysis and design of computer game player experience*, pp. 140-162. Rovaniemi: University of Lapland.

Huotari, K & Hamari, J. (2012). Defining gamification: a service marketing perspective. In *Proceedings of the 16th International Academic MindTrek Conference,* Tampere, Finland, October 3-5, 2012, ACM, pp. 17-22.

Jørgensen, K. (2013). *Gameworld interfaces.* Cambridge, MA: The MIT Press.

Keys, B. & Wolfe, J. (1990). The role of management games and simulations in education and research. *Journal of Management*, 16(2), 307-336.

Kibbee, J. M., Craft, C. J. & Nanus, B. (1961). *Management Games: A new technique for executive development*. New York: Reinhold.

Kim, D. H. (1993). The link between individual and organizational learning. *Sloan Management Review*, 35(1), 37-50.

Klabbers, J. H. G. (2003). Interactive learning of what? In Percival, F., Godfrey, H., Laybourn, P. & Murray, S. (Eds.) *The international simulation & gaming yearbook vol. 11*, pp. 257-266. Edinburgh: Napier University, Edinburgh.

Klabbers, J. H. G. (2009). *The magic circle: Principles of gaming & simulation*, Third and Revised Ed. Rotterdam: Sense Publishers.

Knotts, U. S. Jr. & Keys, J. B. (1997). Teaching strategic management with a business game. *Simulation & Gaming,* 28(4), 377-394.

Lainema, T. (2003). *Enhancing organizational business process perception – experiences from constructing and applying a dynamic business simulation game*. Diss. Turku: Turku School of Economics.

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. New York: Penguin Press.

Meier, R. C., Newell, W. T. & Pazer, H. L. (1969). *Simulation in business and economics*. Englewood Cliffs, NJ: Prentice-Hall.

Myers, D. (2010). *Play redux: The form of computer games.* Ann Arbor, MI: The University of Michigan Press

and the University of Michigan Library.

Palmunen, L.-M., Pelto, E., Paalumäki, A. & Lainema, T. (2013). Formation of novice business students' mental models through simulation gaming. *Simulation & Gaming*, 44(6), 846-868.

Prensky, M. (2001). *Digital game-based learning*. New York: McGraw Hill.

Ritterfeld, L., Cody, M. & Vorderer, P. (Eds.) (2009) *Serious games: Mechanisms and effects*. New York: Routledge.

Tsuchiya, T. & Tsuchiya, S. (1999) The unique contribution of gaming/simulation: Towards establishment of the discipline. In Saunders, D. & Severn, J. (Eds.) *The international simulation & gaming research yearbook: Simulations & games for strategy and policy planning*, pp. 46-57. London: Kogan Page,

Vesa, M. (2013). *There be dragons: An ethnographic inquiry into the strategic practices and process of World of Warcraft gaming groups*. Diss. Helsinki: Hanken School of Economics.

Warmelink, H. (2014). *Online gaming and playful organization*. New York: Routledge.

Whitton, N. (2009). *Learning with digital games: A practical guide to engaging students in higher education*. New York: Routledge.

Whitton, N. (2014). *Digital games and learning: Research and theory*. New York: Routledge.

Zichermann, G. & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps.* Sebastopol: O'Reilly.

Zichermann, G. & Linder, J. (2010). Game-based marketing: Inspiring customer loyalty through rewards, challenges, and contests. Hoboken, NJ: Wiley